ACQUIRING AND USING THREE-DIMENSIONAL INFORMATION IN A DOCUMENT SCANNING SYSTEM

CROSS REFERENCE TO RELATED APPLICATIONS

[0001]   This application is related to co-pending U.S. Patent Application Serial No. _____ entitled "Movable Document Cradle For Facilitating Imaging of Bound Documents" (Attorney Docket No. GOOGP002), U.S. Patent Application Serial No. _____ entitled "Pacing And Error Monitoring Of Manual Page Turning Operator" (Attorney Docket No. GOOGP003), and U.S. Patent Application Serial No. _____ entitled "Imaging Opposing Bound Pages at High Speed Using Multiple Cameras" (Attorney Docket No. GOOGP005), filed concurrently herewith, which are incorporated herein by reference in their entireties.

BACKGROUND OF THE INVENTION

1.   Field of the Invention

[0002]   The present invention relates generally to scanning bound documents. More specifically, systems and methods for acquiring and using three-dimensional imaging data to improve the process of scanning bound documents are disclosed.

2.   Description of Related Art

[0003]   Scanning books, magazines, and other printed material into digital form has become more common with the advent of improved imaging, storage and distribution

techniques. Although unbound printed material can generally be scanned with relative ease using automatic page-feeding mechanisms such as those commonly found on digital copiers and scanners, bound documents present additional challenges. Bound documents include not only books, but also periodicals, manuscripts, pamphlets, brochures, newspapers, manuals, and any other document having a bound edge. Many institutions, such as the Library of Congress, universities, bookstores, and private enterprises have vast collections of bound documents. By converting these documents into electronic form, such institutions can reduce the cost of storage, facilitate remote access, enable simultaneous access by multiple users, facilitate search and retrieval of information, and/or protect information in rare or out-of-print works from loss or destruction.

[0004] Once the content of a bound document is scanned, the recorded image can be manipulated or otherwise processed. Digitally recorded bound documents can be reformatted, supplemented with additional information, compressed, and/or processed with OCR (optical character recognition) software, and indexed to facilitate electronic search. Thus, scanning and recording of bound documents facilitates the creation of digital libraries that can be remotely and simultaneously accessed and searched by multiple users.

[0005] Various mechanisms have been developed to enable the scanning of bound documents. For example, a traditional flat-bed platen scanner scans bound documents in a face-down position. However, for best results, a flat-bed scanner typically requires the application of force to the spine or binding region of the bound documents to insure that they come within the scanner's depth of focus. Such force can damage the spine region of the document. In addition, using the flat-bed platen can be tedious and time-

consuming, as the bound documents typically must be lifted and repositioned after each page is scanned. Further, image quality is often poor due to loss of focus, uneven illumination, and distortion caused by page curvature in the vicinity of the binding.

[0006] An alternative to the traditional flat-bed platen scanner is a platen-less scanner that captures image data from a bound document in a face-up position. Such scanners typically do not require application of additional stress to the binding region of a bound document, since the document is scanned in its natural, face-up position. Some such scanners make use of automatic page turning apparatuses. In addition, optics and software have been developed for compensating for the image-distortion resulting from the inherently curved surfaces typical of open, edge-bound books that are not subjected to potentially damaging flattening pressure. For example, the curvature of a document's pages may be detected using a curvature detector, with a linear light source for defining a brightness distribution. The detected curves of the page can then be used to approximately rectify image signals received from the document scanner.

[0007] A problem with such techniques is that obstructions may appear between the document and the camera or other device that is used to obtain an image of the document. This can reduce the effectiveness and/or throughput of large scanning jobs, since the inaccurate and/or ineffective scanning of one or more pages can undermine the reliability of information search and retrieval based on the results of the scanning process, and renders the scanned version of the document an ineffective and incomplete tool for archival purposes.

[0008]   Accordingly, systems and methods are needed for facilitating the scanning of bound documents in a manner that minimizes document wear-and-tear, and yet is relatively efficient and/or error-free.

5   **SUMMARY OF THE INVENTION**

[0009]   Systems and methods for acquiring and using three-dimensional imaging data to improve the process of scanning bound documents are disclosed. It should be appreciated that the present invention can be implemented in numerous ways, including as a process, an apparatus, a system, a device, a method, or a computer readable medium

10   such as a computer readable storage medium or a computer network wherein program instructions are sent over optical or electronic communication lines. Several inventive embodiments of the present invention are described below.

[0010]   In one embodiment, a document scanning system is provided. The system includes at least one camera for capturing an image of documents such as books,

15   periodicals, and the like, and at least a second camera for gathering data regarding the first camera's field of view. Signal processing logic analyzes the data gathered by the second camera, and uses it to determine whether the imaging camera's view was obstructed. If it is determined that the imaging camera's view was obstructed, then remedial action can be taken, such as removing the obstruction, recapturing the image,

20   and/or flagging the image for further processing. In one embodiment, the second camera may include a laser range finder, a scanning laser range finder, an ultrasonic range finder, a device that utilizes auto-focus range finding technology, a binocular camera array, and/or a trinocular camera array.

[0011]   In another embodiment, a method for scanning bound documents is provided. An array of one or more cameras is used to capture an image of a document. A second array of cameras or other devices are also used to collect data regarding the document and region between the document surface and the first camera array. These data are then processed to determine whether any obstructions or other error conditions might have interfered with the first camera array's attempt to capture an image of the document. If any such obstructions or error conditions are detected, remedial action can be taken. In some embodiments, the process of detecting the presence of obstructions or error conditions includes comparing the data collected by the second array of cameras to an idealized data model of a hypothetical document and the region between the hypothetical document and the first camera array. Differences between the actual data collected by the second array and the idealized model are analyzed to determine whether they indicate the presence of obstructions or other error conditions.

[0012]   In yet another embodiment, a method for scanning a bound document is provided. The document is positioned on a cradle located beneath an array of one or more cameras. The array of cameras is used to capture an image of the bound document. A second array of one or more cameras is used to capture an image of the bound document and the region between the surface of the bound document and the first camera array. The image captured by the second camera array is then compared to data representing a hypothetical, bound document and the region between the surface of the hypothetical, bound document and the first camera array. Differences between the captured data and the data model can then be analyzed to detect the presence of

obstructions or other error-causing conditions, such as warped pages, an incorrectly positioned document, or the like.

[0013]   These and other features and advantages of the present invention will be presented in more detail in the following detailed description and the accompanying

5   figures which illustrate by way of example the principles of the invention.

BRIEF DESCRIPTION OF THE DRAWINGS

[0014]   The present invention will be readily understood by the following detailed description in conjunction with the accompanying drawings, wherein like reference

10   numerals designate like structural elements.

[0015]   **FIG. 1** is a schematic diagram of an exemplary image-capture system.

[0016]   **FIG. 2** is a schematic diagram of an exemplary image-capture system that acquires and uses three-dimensional shape information for error detection.

[0017]   **FIG. 3** is a block diagram illustrating a method for acquiring and using three-

15   dimensional imaging data to improve the process of scanning bound documents.

[0018]   **FIG. 4** is a schematic diagram on another exemplary image-capture system.

[0019]   **FIG. 5** is a schematic diagram of an exemplary image-capture system that uses range-finding technology to detect obstructions.

[0020]   **FIG. 6** is a block diagram of a system for acquiring and using three-

20   dimensional imaging data to improve the process of scanning bound documents.

DESCRIPTION OF SPECIFIC EMBODIMENTS

[0021]   Systems and methods for acquiring and using three-dimensional imaging data to improve the process of scanning bound documents are disclosed. The following description is presented to enable any person skilled in the art to make and use the invention. Descriptions of specific embodiments and applications are provided only as examples, and various modifications will be readily apparent to those skilled in the art. The general principles described herein may be applied to other embodiments and applications without departing from the spirit and scope of the invention. Thus, the present invention is to be accorded the widest possible scope, encompassing numerous alternatives, modifications, and equivalents consistent with the principles and features disclosed herein. For purpose of clarity, technical details that are known in the fields related to the invention have not been described in detail so as not to unnecessarily obscure the present invention.

[0022]   FIG. 1 is a schematic diagram of an exemplary system 20 for scanning bound documents. As shown in FIG. 1, a document such as a book 26 is placed on a cradle 28 (such as a table top, or mechanical device specially adapted for holding books and other documents), and opened to facing pages 30, 32 which a system operator wishes to convert into electronic form. One or more cameras 22, 24, such as high resolution cameras, take digital pictures of pages 30, 32, and these pictures are then processed using optical character recognition (OCR) techniques to produce an editable and/or searchable version of the document's text.

[0023]   In order to scan additional pages of the document 26, it is necessary to open the document 26 to each such page so that an image can be taken. This can be done

manually by the system operator, or automatically using a robot arm or other mechanical mechanism. In embodiments in which the document pages are turned automatically, it will often be desirable for the camera(s) 22, 24 to take pictures automatically at predefined time intervals, synchronized to match the pace of the automatic page-turning mechanism. Even in embodiments where the system operator turns the pages manually, it may still be desirable for the camera(s) to take pictures automatically (e.g., at predefined intervals) in order to increase the speed at which a document can be scanned by obviating the need for additional manual steps (e.g., the step of manually triggering the cameras).

[0024]   When scanning a large document in its entirety (or a substantial portion thereof), it will generally be desirable to process each page as quickly as possible, with relatively little human supervision. It will thus be desirable to automate the process to a high degree, and/or to reduce the level of skill required of any human operators.

[0025]   A problem with a scanning process such as that illustrated in **FIG. 1** is that camera(s) 22, 24 may be inadvertently blocked when taking a picture of a given page, thus resulting in a failure to obtain an accurate and/or complete image of the page. If such an obstructed picture goes undetected, or unremedied, it can lead to error, data loss, and/or customer dissatisfaction when the incompletely and/or inaccurately scanned text is later used as a search tool or as a backup of the original, or if it is marketed to consumers as an accurate and complete copy of the original.

[0026]   It may be possible to reduce such errors by slowing the scanning process down, and/or by relying more heavily on human intervention (or more highly skilled human intervention) to ensure that each page is correctly positioned well in advance of

camera activation, and to ensure that any obstruction to each camera's view is removed. Alternatively, or in addition, the entire text of each document could be reviewed, word-by-word, by human readers once scanning is complete. A problem with such approaches, however, is that they are likely to be time-consuming and expensive, and, due to their

5 reliance on human attentiveness, may nevertheless fail to prevent or detect a substantial number of errors.

[0027] FIG. 2 is a schematic diagram of an exemplary image-capture system that acquires and uses three-dimensional shape information to automatically detect and optionally correct obstructions and/or other problems with the images obtained by

10 camera(s) 22, 24. As shown in FIG. 2, one or more additional cameras 40a, 40b, 40c, 42 may be employed to obtain three-dimensional (or quasi-three-dimensional) spatial data that can be used to detect obstructions in the field of vision of camera(s) 22, 24. The three-dimensional data gathered by cameras 40, 42 may undergo relatively sophisticated signal processing to obtain additional information about the nature of any obstructions, or

15 relatively simple signal processing could be used to make a yes/no decision as to whether any obstructions were present. For example, the signal from camera 42 might simply detect obstructions in camera 42's line of sight to target 44, without undergoing further signal processing. Indeed, such a camera 42 may be a laser beam or ultrasound device. Alternatively or additionally, one or more cameras 42 may be operable to capture an edge

20 view of the document, which can be compared to the edge view of a nominal, idealized document using signal processing techniques, thereby enabling detection of unusual or undesirable profiles. In any event, upon detection of an obstruction or other special condition, the imaging data for a page can be flagged for later remediation, and/or the

human or machine supervisor of the scanning process can be alerted to the error so that the page can be rescanned.

[0028] The imaging data collected by systems such as that shown in **FIG. 2** can also (or alternatively) be used to detect other errors or special conditions during the imaging process. For example, the imaging data collected by cameras 40, 42 can be used to determine whether the document, and/or page, is positioned correctly relative to the image plane(s) of the camera(s) 22, 24; whether the page contains any special features (such as pop-up text) requiring special attention; whether the end of the document has been reached (e.g., by detecting that the document is shut); and/or the like.

[0029] It should thus be appreciated that **FIG. 2** is provided for purposes of illustration, and that numerous modifications could be made to the system shown in **FIG. 2** without departing from the spirit or scope of the present invention. For example, although **FIG. 2** shows one set of cameras 40, 42 being used for purposes of detecting obstructions to the field of view of another set of cameras 22, 24, it will be appreciated that in other embodiments some or all of these cameras may perform both roles (i.e., obtaining imaging data and assisting in the detection of obstructions or other special conditions). In addition, although **FIG. 2** shows two cameras being used for imaging, and four cameras being used to gather three-dimensional data, it should be appreciated that any suitable number of cameras can be used for the image collection and/or three-dimensional data gathering. For example, a single camera could be employed to image the document, and a single camera could be used to detect obstructions. Generally speaking, however, the greater the number of cameras used to collect the three-

dimensional data, the richer the information regarding obstructions and other issues that will be obtainable.

[0030] FIG. 3 illustrates a process for using three-dimensional imaging data to detect and optionally correct errors and/or special conditions in a document imaging process. Note that although for ease of explanation the imaging data that is used for detecting errors and special conditions will often be referred to as "three-dimensional" data, this data need not be three-dimensional in a literal sense. For example, in a simple embodiment, data from a single camera or other data collection source could comprise the "three-dimensional" data that is used.

[0031] Referring to FIG. 3, a document is scanned by positioning it under the imaging camera(s) at step 300. An image of the document is then captured, as is three-dimensional imaging data regarding the space surrounding the document at step 302. The three-dimensional data is then sent to a signal processor for processing at step 304. If the signal processing detects an obstruction and/or other special condition (i.e., a "Yes" exit from decision block 306), then the imaging data for that page is flagged for further processing and/or corrected at step 308. For example, depending on the nature of the condition that is detected, the image data for that particular page could be recaptured, or reprocessed using special signal processing techniques, or simply flagged for manual proofreading. Certain automated remedial measures can be employed at step 308 such as by using software to modify the captured image data and/or by physically correcting the source of the problem.

[0032] Referring once again to **FIG. 3**, if an error or other condition is not detected (i.e., a "No" exit from decision block 306), then the next page can be scanned, the process repeating itself until no more pages remain to be imaged.

[0033] An advantage of the process shown in **FIG. 3** is that potential errors are

5 detected as they occur, thereby allowing immediate remediation. For example, a human or machine operator could be alerted at step 308 to simply recapture the image for the particular page, thereby obviating the need to later return to that page to capture a better image (if the error is detected at all). The spatial data processing may also alert the operator of the need to reposition the document and/or the imaging cameras, thereby

10 preventing the collection of an entire document's worth of useless data due to a careless initial placement of the document under the imaging cameras. Alternatively, some or all of the errors or special conditions detected at step 306 could be processed at a later time. For example, if it is determined that a page contains a pop-up image, the scanning system may return to that page at a later time so that it can apply specialized image capture

15 techniques (e.g., additional, or repositioned, cameras). In sum, the automatic error detection provided by embodiments of the present invention facilitates efficient quality control over the imaging process.

[0034] As previously indicated, any of a wide variety of techniques can be used to collect the three-dimensional data used to detect obstructions. One such technique makes

20 use of a set of one or more cameras or lasers oriented in one or more planes parallel to the face of the document. **FIG. 4** shows an example of one such arrangement. As shown in **FIG. 4**, detection devices such as cameras or lasers 42a…42z are operable to detect obstructions to their field of view. By using a matrix of such devices, the document

scanning system can obtain detailed information about the nature and location of any obstruction to the face of document 26. In some embodiments, multiple matrices of detection devices can be stacked vertically in relation to the page surface, and/or the detection devices in a single matrix can be distributed at different heights, thereby

5   enabling detection of obstructions anywhere between the page surface and the image capturing cameras. In other embodiments, scanning lasers could be used instead of, or in addition to, a static array of lasers such as that shown in **FIG. 4**.

[0035]   In other embodiments, other mechanisms could be used to gather three-dimensional information. For example, incomplete three-dimensional information

10   obtained by, e.g., one or more ultrasonic, laser, and/or auto-focus range finding devices can be combined with a model of an idealized document (e.g., book) to determine whether potential obstructions are present between the document's surface and the imaging camera(s). For example, the scanning system can determine whether the distance to the closest object detected by the range finding device(s) is less than the

15   distance to the hypothesized surface of a properly positioned model book. The detection of an object within the predefined range could indicate a potential error condition that might comprise the effectiveness of the imaging process. Such a technique need not involve a high degree of signal processing—simply examining the output of the range finder could suffice—yet can be quite effective in detecting a wide-range of obstructions

20   or other error conditions, such as an operator's hand, a page that has not settled into its proper position, an oversized document, a wrinkled or distorted page, and the like.

[0036]   An example of such an arrangement is shown in **FIG. 5**. A model of an idealized book 70 is used by the document scanning system's signal processing logic to

interpret the data obtained by scanning range-finding laser 80. Laser 80 preferably collects its data at approximately the same time as imaging camera 82 captures the image of an actual document 72. A potential error condition may be detected, for example, if the distance detected by the range-finding laser 80 at a particular point is less than the

5   distance from the laser to the outline of the idealized book 70 at that point. In **FIG. 5**, for example, the operator's hand 78 blocks laser beam 60 from reaching the document 72. The system detects this obstruction by comparing the length of ray 60 to the distance between laser 80 and the idealized book 70 at that angle. The obstruction caused by the incompletely turned page at ray 64 would be detected in the same manner. Rays 62 and

10 66, on the other hand, reach the surface of the actual document 72. Because the length of each of these rays is greater than the corresponding distance between the laser 80 and the idealized book 70, no obstruction is deemed to exist between the surface of book 72 and imaging camera 82 at these angles.

[0037]   Thus, it should be appreciated that any of a variety of techniques (or

15 combinations thereof) could be used to gather the spatial data used to detect error conditions during the imaging process, non-limiting examples of which include, stereo, trinocular, and other multi-camera arrangements; scanning laser range finders; simple laser range finders; ultrasonic range finders; auto-focus range determination technology; sparse point maps; and/or the like.

20 [0038]   **FIG. 6** shows a detailed block diagram of an illustrative document scanning system. As shown in **FIG. 6**, scanning system 600 may include a processor 602, system memory 604, one or more cameras 610 for capturing document images, one or more devices 612 for collecting three-dimensional spatial information, a mechanism 614 (e.g.,

a flash, array of lamps, etc.) for controlling the lighting of the document surface, an apparatus 616 for turning document pages automatically, and a bus 618 for connecting the aforementioned elements.

[0039]   In a preferred embodiment, the operation of system 600 is controlled primarily by processor 602 operating under the guidance of programs stored in memory 604. Memory 604 may include both high-speed random-access memory (RAM) and non-volatile memory such as a magnetic disk, flash EEPROM, and/or the like.

[0040]   As shown in **FIG. 6**, memory 604 may include a variety of programs or modules for controlling the operation of system 600 and performing the techniques described above in connection with **FIGS. 1-5**. For example, memory 604 may include a collection of images 629 received from cameras 610, and/or data that has been generated from such images using optical character recognition software 630. Memory 604 also preferably includes a signal processing program 631 for analyzing the three-dimensional data gathered by devices 612 to detect error-conditions or other special conditions. For example, signal processing program 631 may be configured to analyze the gathered data with reference to a nominal three-dimensional model of a properly positioned document, with data that does not conform to the model indicating the presence of an obstruction or other special situation. Memory 604 may also include an operating system 632 for coordinating the operation of the various programs and/or hardware, and for taking action in response to the detection of errors or other special conditions.

[0041]   One of ordinary skill in the art will appreciate that the systems and methods of the present invention can be practiced with devices and architectures that lack many of the components shown in **FIG. 6** and/or that have other components that are not shown.

For example, some systems may not include an automatic page turning apparatus 616, a separate group of devices 612 for performing three-dimensional data capture, an automatically controlled lighting mechanism 614, and/or OCR software 630. Other systems may employ mechanisms that are not shown in **FIG. 6**, such as special-purpose

5   signal processing hardware (e.g., one or more digital signal processing (DSP) chips), and/or the like. In other embodiments, the functionality shown in **FIG. 6** could be distributed amongst multiple systems (e.g., image data could be sent to a remote location for application of OCR, indexing, and storage). Thus, it should be appreciated that **FIG. 6** is provided for purposes of illustration and not limitation.

10   **[0042]**   As previously indicated, spatial information collected in the manner described above can be used to detect a wide variety of error conditions and special situations. Several examples have been given above. Additional examples include the detection and optional correction of page distortions caused by clamps or other mechanisms used to hold a document in place; detection and optional correction of pages outside the focus

15   range of the imaging camera(s); detection and optional correction of document misalignment on the cradle, detection and optional correction of books that are too large or too small; and detection and optional correction of unusual document surfaces, such as pop-ups or foldouts, to name but a few non-limiting examples. Depending on the nature of the obstruction and/or error that is detected, in some embodiments certain automated

20   remedial measures can be employed such as by using software to modify the captured image data and/or by physically correcting the source of the problem. Examples of automated remedial measures include physically adjusting the lighting and/or using software to compensate for shadowing such as may be caused by page warp,

mechanically flattening and/or using software to compensate for warped pages, ensuring that recently turned pages have completely settled, adjusting the optical character recognition (OCR) process to compensate for distortions caused by poorly positioned documents or defective document surfaces, controlling the focus of the imaging camera to compensate for abnormally large or small documents, controlling the position of a motorized document holding cradle, determining when to take special actions like firing a high-definition camera or flash, and/or driving a statistically controlled quality control process by identifying "troubled" pages.

[0043]   While the preferred embodiments of the present invention are described and illustrated herein, it will be appreciated that they are merely illustrative and that modifications can be made to these embodiments without departing from the spirit and scope of the invention. Thus, the invention is intended to be defined only in terms of the following claims.